



T.C.

İSTANBUL TOPKAPI

ÜNİVERSİTESİ

MÜHENDİSLİK FAKÜLTESİ

YAZILIM MÜHENDİSLİĞİ BÖLÜMÜ

AKCİĞER KANSERİ RİSK TAHMİNİ UYGULAMASI

SELİN ŞENOCAK

21040301043

BİTİRME PROJESİ

İSTANBUL 2025

T.C.
İSTANBUL TOPKAPI
ÜNİVERSİTESİ
MÜHENDİSLİK FAKÜLTESİ
YAZILIM MÜHENDİSLİĞİ BÖLÜMÜ

AKCİĞER KANSERİ RİSK TAHMİNİ UYGULAMASI

SELİN ŞENOCAK
21040301043

Projenin Danışmanı: Doç. Dr. Cevat RAHEBİ

AKADEMİK DÜRÜSTLÜK BEYANI

Mezuniyet projesi olarak sunduğum “Akciğer Kanseri Risk Tahmin Uygulaması” başlıklı çalışmanın, bilimsel ahlak ve geleneklere uygun olarak tarafımdan yazıldığını, yararlandığım eserlerin tamamının kaynaklarda gösterildiğini ve çalışmam içerisinde kullanıldıkları her yerde atıf yapıldığını belirtir ve onurumla doğrularım.

19.06.2025

SELİN ŞENOCAK

TEŞEKKÜR

Çalışmalarım süresince bilimsel katkıları ile beni aydınlatan, farklı bakış açıları ile yol gösteren, yakın ilgi ve yardımları ile beni motive eden ve bu günlere gelmemde en büyük katkı sahibi olan sayın Doç. Dr. Cevat Rahebi'ye teşekkürü bir borç bilirim.

Ayrıca, çalışmalarım süresince sabır göstererek beni daima destekleyen değerli aileme ve arkadaşlarıma en içten teşekkürlerimi sunarım.

19.06.2025

SELİN ŞENOCAK

ÖZET

Akciğer kanseri, hem erkeklerde hem de kadınlarda dünyanın en çok ölüme neden olan kanser türlerinden biri olmasıyla dikkat çekmektedir. Dünyada her yıl milyonlarca kişiye bu hastalık tanısı konulmakta ve önemli bir bölümü, hastalığın geç evrelerinde fark edildiği için tedaviye yanıt verememektedir. Bu bağlamda erken tanı, hastalığın seyri açısından kritik bir rol oynamakta, bireylerin hayatlarını kurtarabilecek önemli bir adım olarak karşımıza çıkmaktadır. Geleneksel tanı yöntemlerinin geçerliliği sürmekle birlikte, dijital sağlık teknolojilerinin gelişmesiyle birlikte erken uyarı sistemlerinin dijital ortamda sunulması, hem maliyeti azaltmakta hem de toplumsal farkındalığı artırmaktadır.

Bu bitirme projesinde, kullanıcıdan alınan sağlık ve yaşam tarzı verileri ile akciğer kanseri riskini tahmin edebilen, mobil tabanlı, yapay zeka destekli bir uygulama geliştirilmiştir. Sistem Flutter kullanılarak geliştirilen bir mobil uygulamadan; Flask ile yazılmış bir backend API katmanından; Scikit-learn kütüphanesiyle eğitilmiş bir makine öğrenmesi modelinden ve verilerin saklandığı MSSQL veritabanından oluşmaktadır. Kullanıcı, uygulama üzerinden yaş, sigara kullanımı, yorgunluk, anksiyete gibi 15+ parametreyi doldurduktan sonra sistem bu verileri Flask API aracılığıyla backend'e göndermektedir. Gelen veriler, önceden SMOTE uygulanarak dengelenmiş ve StandardScaler ile normalize edilmiş bir veri seti üzerinde Logistic Regression modeliyle eğitilmiş tahmin modeline aktarılmakta; tahmin sonucu hem ekranda görüntülenmekte hem de MSSQL veritabanına zaman damgasıyla birlikte kayıt altına alınmaktadır.

Uygulama sadece bir tahmin motoru olmaktan çıkmış; aynı zamanda kullanıcıya geçmiş tahminlerini de sunabilen bir sisteme dönüşmüştür. "Sonuçları Gör" ekranında kullanıcı, hem kendi son tahminini hem de son iki test sonucunu özet halinde inceleyebilmektedir.

Makine öğrenmesi modelinin eğitimi sürecinde veri ön işleme, encoding, SMOTE ile dengeleme, ölçekleme ve modellerin karşılaştırılması gibi adımlar uygulanmış; Logistic Regression modeli diğer modellerle karşılaştırıldığında en dengeli performansı (f1-score, recall) sunması sebebiyle tercih edilmiştir. Eğitilen model .pkl formatında saklanmış ve Flask uygulamasına entegre edilmiştir.

Bu proje, sađlık teknolojilerinin yapay zeka ile buluřtuđu noktada farkındalık yaratmayı, erken uyarı mekanizmalarını herkesin cebine taşımayı ve tıp verilerinin sadece bir tahlil sonucundan ibaret olmadığını göstermeyi hedeflemektedir. Bilimsel doğruluđun, yazılımsal kaliteyle buluřtuđu bu uygulama; hem bireylerin kendi sađlıklarını gözetmelerini hem de toplumsal bazda erken teşhis imkanlarının artmasını desteklemektedir. Uygulama, gelecekte uzman görüşü entegrasyonu, veriye dayalı öğrenme sistemi ve doktorlar için panel gibi modüllerle geliştirilebilir bir altyapı sunmaktadır.

ABSTRACT

Lung cancer is one of the most fatal types of cancer affecting both men and women worldwide. Millions of individuals are diagnosed with this disease every year, and a significant portion of them cannot respond to treatment due to late-stage detection. In this context, early diagnosis plays a critical role in the course of the disease and stands out as a vital step in saving lives. While traditional diagnostic methods remain valid, the development of digital health technologies enables the delivery of early warning systems in a digital environment, reducing costs and increasing public awareness.

In this graduation project, a mobile-based, AI-powered application has been developed to predict lung cancer risk using health and lifestyle data collected from users. The system consists of a Flutter-based mobile application, a Flask-based backend API, a machine learning model trained with the Scikit-learn library, and an MSSQL database for data storage. The user fills out a form with over 15 parameters, including age, smoking habits, fatigue, and anxiety, and this data is sent to the backend through the API. The incoming data is processed using a previously SMOTE-balanced and StandardScaler-normalized dataset, and the prediction is made using a Logistic Regression model. The result is displayed on the screen and simultaneously stored in the MSSQL database with a timestamp.

The application has evolved from being a mere prediction engine into a system that also provides users access to their previous results. On the "View Results" screen, users can review both their most recent prediction and a summary of their last two tests.

During the training process of the machine learning model, several steps were applied, including data preprocessing, encoding, balancing with SMOTE, scaling, and model comparison. Logistic Regression was selected for deployment due to its balanced performance in metrics such as f1-score and recall. The trained model was saved in .pkl format and integrated into the Flask backend.

This project aims to raise awareness at the intersection of artificial intelligence and healthcare technologies, make early warning systems accessible to everyone, and demonstrate that medical data is more than just lab test results. By combining scientific accuracy with software quality,

this application supports both personal health monitoring and broader societal access to early diagnosis opportunities. It also offers a flexible infrastructure for future modules such as expert opinion integration, adaptive learning systems, and a doctor-specific panel.

İçindekiler

AKADEMİK DÜRÜSTLÜK BEYANI.....	ii
TEŞEKKÜR.....	iii
ÖZET.....	iv
ABSTRACT.....	vi
Şekiller Dizini.....	x
Çizelgeler Dizini.....	xi
1. GİRİŞ.....	1
1.1 Projenin Amacı ve Önemi	1
1.2 Sistem ve Çalışma Prensibi.....	2
1.3 Projenin Hedefleri ve Beklenen Çıktılar	3
2. LİTERATÜR TARAMASI.....	5
2.1 Akciğer Kanseri Tahmini Üzerine Yapılan Çalışmalar.....	5
2.2 Yapay Zekâ Destekli Sağlık Uygulamaları.....	6
2.3 Mobil Tabanlı Teşhis Sistemleri.....	7
3. SİSTEM ANALİZİ	9
3.1 Problem Tanımı ve Motivasyon	9
3.2 Sistem Gereksinimleri.....	11
3.2.1 İşlevsel Gereksinimler	11
3.2.2 İşlevsel Olmayan Gereksinimler.....	11
3.3 Hedef Kitle ve Kullanım Senaryoları.....	12
3.3.1 Hedef Kullanıcı Profilleri.....	12
3.3.2 Örnek Kullanım Senaryoları.....	12
4. VERİ SETİ VE ÖN İŞLEME.....	14
4.1 Kullanılan Veri Seti.....	14
4.2 Veri Temizleme Süreci.....	15
4.3 Veri Dengeleme ve SMOTE Uygulaması	15
4.4 Ölçekleme (Scaling).....	16
5. MODEL EĞİTİMİ VE PERFORMANS DEĞERLENDİRME	18
5.1 Model Seçimi Süreci	18
5.2 Değerlendirme Metrikleri	18
5.3 Performans Karşılaştırması	18
5.4 Modelin Kaydedilmesi ve Entegrasyonu	20
6. UYGULAMA GELİŞTİRME VE MİMARİ YAPI	21

6.1 Genel Mimari Yapı	21
6.2 Kullanılan Teknolojiler.....	26
6.3 Modüler Kod Yapısı	26
6.4 Uçtan Uca Veri Akışı.....	27
7. TASARIM VE DENEYSEL DEĞERLENDİRME.....	29
7.1 Arayüz Tasarım Prensipleri.....	29
7.2 Performans Değerlendirmesi (Planlama).....	29
7.3 Kullanıcı Testleri ve Geri Bildirimler (Planlama).....	30
8. TARTIŞMA VE SONUÇ	31
8.1 Elde Edilen Sonuçların Yorumlanması	31
8.2 Karşılaşılan Zorluklar ve Öğrenilenler	31
8.3 Projenin Literatüre ve Uygulamaya Katkısı.....	32
8.4 Gelecek Çalışmalar ve İyileştirme Önerileri.....	32
9. PLANLAMA	33
9.1 Proje Zaman Çizelgesi ve Aşamaları	33
9.2 Kaynak Yönetimi.....	34
9.3 Risk Yönetimi.....	34
10. KAYNAKÇA	35

Şekiller Dizini

Şekil 1: Kullanıcı ile sistem arasındaki temel kullanım senaryoları (use case diyagramı).....	16
Şekil 2:Akciğer kanseri testi mobil uygulama aktivite diyagramı.....	17
Şekil 3:Makine öğrenmesi modellerinin Accuracy, Recall ve F1-Score metriklerine göre görsel karşılaştırması.....	19
Şekil 4:Akciğer kanseri testi için kullanıcıdan veri alınan test formu arayüzü	22
Şekil 5:Kullanıcının tahmin sonucunun “Düşük risk” olarak gösterildiği ekran.....	23
Şekil 6:Kullanıcının tahmin sonucunun “Yüksek risk” olarak gösterildiği ekran.....	24
Şekil 7:Uygulama ana ekranı: test başlatma ve geçmiş sonuçları görüntüleme seçenekleri	25
Şekil 8:Mobil uygulama, API ve sunucu arasındaki sekans diyagramı.....	28

Çizelgeler Dizini

Tablo 1. Veri setinde yer alan değişkenler ve açıklamaları.....	14
Tablo 2. Modellerin sınıflandırma performans karşılaştırması.....	19
Tablo 3. İş Gücü Hesaplama Tablosu	33
Tablo 4. Risk Yönetimi.....	34

1. GİRİŞ

Sağlık alanında dijital dönüşümün hız kazanmasıyla birlikte, yapay zekâ destekli erken teşhis sistemleri giderek daha önemli hale gelmektedir. Özellikle ölüm oranı yüksek olan hastalıklarda, risk tahminine dayalı dijital araçlar sayesinde bireylerin sağlık durumlarına ilişkin erken önlem alabilmeleri mümkün hale gelmiştir. Bu bağlamda geliştirilen bu bitirme projesi, akciğer kanseri riskinin belirli parametrelere dayalı olarak tahmin edilmesini amaçlayan mobil tabanlı bir uygulamayı kapsamaktadır. Uygulama, kullanıcıların doldurduğu sağlık ve yaşam tarzı verilerini kullanarak makine öğrenmesi tabanlı bir model yardımıyla tahmin yapmakta ve sonucu hem kullanıcıya sunmakta hem de veritabanına kaydetmektedir.

Geliştirilen sistem, mobil uygulama (Flutter), sunucu taraflı servis (Flask), makine öğrenmesi modeli (Scikit-learn) ve veritabanı katmanı (MSSQL) olmak üzere dört ana bileşenden oluşmaktadır. Kullanıcı tarafından sağlanan bilgiler, Flask API aracılığıyla sunucuya gönderilmekte ve burada işlenerek eğitilmiş model tarafından risk analizi yapılmaktadır. Bu sonuçlar hem uygulama arayüzünde görüntülenmekte hem de MSSQL veritabanına zaman damgası ile birlikte kaydedilmektedir. Ayrıca uygulama, geçmiş test sonuçlarını da kullanıcıya sunarak hem kişisel takip hem de toplu analiz yapılabilmesini sağlamaktadır.

Bu bölümde, projenin amacı, içeriği, genel çalışma prensibi ve uygulama hedefleri ayrıntılı biçimde ele alınacaktır. Ayrıca sistemin sağlayacağı faydalar, katkılar ve hedeflenen çıktılar da açıklanarak sonraki bölümlere temel oluşturulacaktır.

1.1 Projenin Amacı ve Önemi

Sağlık alanında dijital dönüşümün hız kazanmasıyla birlikte, yapay zekâ destekli erken teşhis sistemleri giderek daha önemli hale gelmektedir. Özellikle ölüm oranı yüksek olan hastalıklarda, risk tahminine dayalı dijital araçlar sayesinde bireylerin sağlık durumlarına ilişkin erken önlem alabilmeleri mümkün hale gelmiştir. Bu bağlamda geliştirilen bu bitirme projesi, akciğer kanseri riskinin belirli parametrelere dayalı olarak tahmin edilmesini amaçlayan mobil tabanlı bir uygulamayı kapsamaktadır. Uygulama, kullanıcıların doldurduğu sağlık ve yaşam tarzı verilerini kullanarak makine öğrenmesi tabanlı bir model yardımıyla tahmin yapmakta ve sonucu hem kullanıcıya sunmakta hem de veritabanına kaydetmektedir.

Geliştirilen sistem, mobil uygulama (Flutter), sunucu taraflı servis (Flask), makine öğrenmesi modeli (Scikit-learn) ve veritabanı katmanı (MSSQL) olmak üzere dört ana bileşenden oluşmaktadır. Kullanıcı tarafından sağlanan bilgiler, Flask API aracılığıyla sunucuya gönderilmekte ve burada işlenerek eğitilmiş model tarafından risk analizi yapılmaktadır. Bu sonuçlar hem uygulama arayüzünde görüntülenmekte hem de MSSQL veritabanına zaman damgası ile birlikte kaydedilmektedir. Ayrıca uygulama, geçmiş test sonuçlarını da kullanıcıya sunarak hem kişisel takip hem de toplu analiz yapılabilmesini sağlamaktadır.

Bu bölümde, projenin amacı, içeriği, genel çalışma prensibi ve uygulama hedefleri ayrıntılı biçimde ele alınacaktır. Ayrıca sistemin sağlayacağı faydalar, katkılar ve hedeflenen çıktılar da açıklanarak sonraki bölümlere temel oluşturulacaktır.

Projenin temel amacı, bireylerin sağlık durumlarını basit ve anlaşılır bir arayüz üzerinden değerlendirerek akciğer kanseri riskini erken aşamada öngörmelerine olanak sağlamaktır. Bu doğrultuda geliştirilen mobil uygulama, kullanıcının yaş, sigara alışkanlığı, solunum durumu gibi sağlık ve yaşam tarzı bilgilerinden yola çıkarak bir risk tahmini üretmektedir. Uygulamanın temel özelliği, tahmini yalnızca kullanıcıya sunmakla kalmayıp, aynı zamanda bu verileri düzenli olarak veritabanına kaydederek ileride yapılacak toplu analizlere de katkı sağlamasıdır.

Proje, düşük maliyetli bir erken teşhis destek sistemi geliştirilmesi açısından büyük önem taşımaktadır. Özellikle sağlık hizmetlerine erişimi kısıtlı bireyler için bu uygulama, risk farkındalığını artırmakta ve erken müdahale için bir uyarı mekanizması olarak görev yapmaktadır. Ayrıca geliştirilen altyapı, sağlık alanında dijital veri analitiği ve kullanıcı odaklı yazılım sistemlerinin bir araya getirildiği örnek bir model sunmaktadır.

1.2 Sistem ve Çalışma Prensibi

Bu projede geliştirilen sistem, kullanıcıdan alınan verilerin bir dizi işlemten geçerek yapay zeka tabanlı bir modelle analiz edilmesini ve sonucun hem kullanıcıya sunulmasını hem de veritabanına kaydedilmesini sağlayan bir yapıyı içermektedir. Sistem Flutter mobil uygulaması üzerinden çalışmaktadır ve bu uygulama aracılığıyla kullanıcı, kendisine yöneltilen sorulara cevap verir. Bu cevaplar, JSON formatında Flask API'sine iletilir. Flask tarafında, bu

veriler önceden SMOTE ile dengelenmiş ve StandardScaler ile normalize edilmiş veri seti üzerinde eğitilmiş Logistic Regression modeliyle analiz edilir.

API katmanı, gelen verileri .pkl formatında saklanan modele aktararak tahmin sonucunu elde eder. Bu sonuç hem mobil uygulamaya yanıt olarak geri dönlmekte hem de MSSQL veritabanına tarih ve saat bilgisiyle birlikte kayıt edilmektedir. Veritabanı, sistemin sadece bir tahmin motoru olmasının ötesinde; toplu analiz, istatistiksel grafik üretimi ve sistematik değerlendirme gibi çok yönlü kullanımlara imkan tanır niteliktedir.

Sistemin modüler yapısı aşağıdaki gibi özetlenebilir:

- **Mobil Uygulama (Flutter):** Kullanıcıdan veri alma, sonucu gösterme.
- **Backend API (Flask):** Verileri alma, modeli çağırma, veritabanına yazma.
- **Makine Öğrenmesi Modeli:** Logistic Regression ile tahmin.
- **Veritabanı (MSSQL):** Girdi-yanıt saklama, analiz verisi biriktirme.

Bu yapı sayesinde sistem; kullanıcı dostu, veri odaklı ve geliştirilebilir bir erken uyarı platformu haline gelmektedir.

1.3 Projenin Hedefleri ve Beklenen Çıktılar

Bu projenin genel hedefi, dijital sağlık teknolojileri ile makine öğrenmesini birleştirerek bireylerin akciğer kanseri riski konusunda erken aşamada bilgilendirilebileceği ve potansiyel olarak farkındalığı artıracak bir sistem ortaya koymaktır. Proje bu hedefe, mobil cihazlardan erişilebilen, kullanıcı dostu bir arayüz ile sağlık ve yaşam tarzı bilgilerini toplayan ve bu verileri yapay zekâ destekli analiz ile işleyerek anlamlı bir tahmin sunan bir uygulama geliştirerek ulaşmayı amaçlamaktadır.

Belirli hedefler şu şekilde sıralanabilir:

- Kullanıcıdan sağlık ve yaşam tarzı verilerinin toplanması (yaş, sigara, öksürük vb.).
- Bu verilerin Flask tabanlı bir API aracılığıyla backend'e iletilmesi.
- Logistic Regression modeli ile verilerden kanser riski tahmini yapılması.
- Tahmin sonucunun mobil uygulamada görüntülenmesi.

- Kullanıcı verisinin MSSQL veritabanına kaydedilmesi.
- Kullanıcıya son iki tahmini özetleyen ekran tasarımının uygulanması.
- Model doğruluğunun istatistiksel olarak test edilmesi (f1, recall, accuracy).

Beklenen çıktılar:

- Flutter destekli çalışan ve test formunu barındıran bir mobil uygulama.
- Flask API ile çalışan tahmin ve kayıt sistemi.
- Logistic Regression modeline dayalı tahmin doğruluk raporu.
- Veritabanına işlenmiş test verileri.
- Kullanıcı görünümleri ile desteklenmiş sonuç ekranları.
- Rapor içinde sunulan şekil, tablo ve grafiklerle desteklenen analiz içerikleri.

Projenin bu hedef ve çıktılarına ulaşması, hem bireysel sağlık farkındalığını artıracak hem de dijital sağlık alanında etkili ve uygulanabilir bir örnek oluşturacaktır.

2. LİTERATÜR TARAMASI

Sağlık teknolojilerinde yapay zekâ uygulamaları son yıllarda hızla artış göstermekte ve özellikle erken teşhis alanında önemli rol oynamaktadır. Literatürde, akciğer kanseri gibi ölüm oranı yüksek hastalıkların risk tahmini üzerine yapılan çalışmalar, hem görüntü işleme hem de yapılandırılmış veri analizi temelli modelleri içermektedir. Bu projede tercih edilen yöntem, kullanıcıdan doğrudan alınan anket temelli sağlık ve yaşam tarzı verileriyle tahmin üretmeye dayandığından; benzer yapıdaki modellerin ve yöntemlerin incelenmesi proje için rehberlik edici olmuştur.

Bu bölümde, akciğer kanseri tahminine yönelik veri odaklı uygulamalardan, mobil tabanlı sağlık uygulamalarına; makine öğrenmesi ile teşhis sistemlerinden kullanıcı etkileşimli risk değerlendirme sistemlerine kadar farklı başlıklarda literatürdeki yaklaşımlara yer verilmektedir. Ayrıca, kullanılan modellerin doğruluk düzeyleri, veri kaynakları ve başarı kriterleri bakımından karşılaştırmalar da yapılacaktır.

Alt bölümlerde, öncelikle sağlık alanında yapay zekâ kullanımı ve ardından akciğer kanseri odaklı çalışmalar incelenmekte; bu proje ile benzerlik ve farklılık gösteren yönler vurgulanmaktadır.

2.1 Akciğer Kanseri Tahmini Üzerine Yapılan Çalışmalar

Akciğer kanseri tahmini üzerine yapılan akademik çalışmalar genellikle iki temel yaklaşımı içermektedir: görüntü tabanlı analizler ve semptom/veri tabanlı analizler. Görüntü tabanlı analizlerde bilgisayarlı tomografi (BT) veya röntgen gibi tıbbi görüntüler üzerinden derin öğrenme algoritmalarıyla tümör tanısı konulması hedeflenmektedir. Ancak bu yaklaşım genellikle uzmanlık gerektiren donanım altyapısı ile sınırlıdır ve mobil tabanlı uygulamalara doğrudan entegre edilmesi zordur.

Buna karşın, yapılandırılmış veri tabanlı modellerde; bireylerin sağlık geçmişi, yaşam tarzı (sigara kullanımı, alkol, yaş, cinsiyet vb.), semptomları (öksürük, göğüs ağrısı, nefes darlığı gibi) ve diğer klinik bilgiler doğrultusunda risk analizi yapılmaktadır. Bu kapsamda yapılan bazı çalışmalarda Logistic Regression, Random Forest, Naive Bayes ve Support Vector Machine (SVM) gibi makine öğrenmesi algoritmalarının karşılaştırmalı olarak kullanıldığı görülmektedir. Elde edilen sonuçlar, Logistic Regression gibi daha yorumlanabilir modellerin,

küçük ve dengelenmiş veri setlerinde başarılı tahmin performansı gösterebildiğini ortaya koymaktadır.

Bazı literatür örneklerinde, özellikle Hindistan ve Çin gibi ülkelerde halk sağlığı taramaları için anket bazlı uygulamaların yaygın şekilde kullanıldığı ve bu yöntemle oluşturulan yapay zekâ sistemlerinin %80'in üzerinde doğruluk sağladığı belirtilmiştir. Bu bağlamda, kullanıcıdan doğrudan alınan anket verileriyle çalışan sistemler, maliyet etkinliği ve erişilebilirlik açısından oldukça avantajlı görünmektedir [1].

Bu projede kullanılan yöntem, semptom ve yaşam tarzı temelli girişleri kullanarak Logistic Regression modeli ile tahmin yapılması üzerine kuruludur ve yukarıda açıklanan çalışmalara metodolojik olarak büyük ölçüde benzerlik göstermektedir. Ancak bu projeyi farklılaştıran yön, sistemin mobil bir arayüz üzerinden çalışması, tahmin sonuçlarını anlık olarak kullanıcıya sunması ve aynı zamanda veri kaydı yaparak uzun vadeli analizlere açık bir yapı sunmasıdır.

2.2 Yapay Zekâ Destekli Sağlık Uygulamaları

Günümüzde sağlık alanında yapay zekâ (YZ) uygulamaları yalnızca klinik tanı ve tedavi süreçlerinde değil, aynı zamanda halk sağlığı, erken teşhis, ilaç geliştirme ve bireysel sağlık takibi gibi birçok farklı alanda kullanılmaktadır. Gelişmiş ülkelerde YZ'nin kullanımı genellikle büyük veri kümeleri ve görüntüleme sistemleri ile sınırlı kalırken; mobil uygulamalar ve kullanıcı temelli sistemler, gelişmekte olan ülkelerde daha yaygın ve erişilebilir çözümler sunmaktadır.

Özellikle mobil sağlık (mHealth) uygulamaları, bireylerin kendi sağlık durumlarını değerlendirebilmesi için kolay ulaşılabilir araçlar sağlamaktadır. Bu tür uygulamalarda YZ algoritmaları; semptom analizi, kronik hastalık takibi, ilaç etkileşim denetimi ve risk skorumlama gibi amaçlarla entegre edilmektedir. Kullanıcıdan alınan basit verilerin (örneğin yaş, cinsiyet, sigara kullanımı, semptomlar) yorumlanması yoluyla sağlık profesyonellerine ulaşmadan önce bir ön değerlendirme yapılabilmektedir.

Yapılan bazı çalışmalarda, özellikle solunum yolu hastalıkları, diyabet, hipertansiyon gibi yaygın rahatsızlıkların takibinde karar destek sistemlerinin mobil cihazlar ile bütünleştirildiği görülmektedir. Örneğin, WHO tarafından desteklenen bazı projelerde,

gelişmekte olan ülkelerde kullanılan anket bazlı mobil uygulamaların, halk sağlığı taramaları için önemli çıktılar ürettiği belirtilmektedir [2].

Bu bağlamda geliştirilen Flutter tabanlı mobil uygulama ve Flask API ile çalışan bu proje; kullanıcı etkileşimli bir sistem olarak, literatürdeki bu eğilimle örtüşmekte ve bireylerin sağlık verilerine dayalı bilinçli kararlar almasına yardımcı olabilecek bir araç olarak konumlandırılmaktadır.

2.3 Mobil Tabanlı Teşhis Sistemleri

Mobil cihazların yaygınlaşmasıyla birlikte, teşhis ve sağlık izleme süreçleri giderek taşınabilir ve kullanıcı merkezli hale gelmektedir. Mobil tabanlı teşhis sistemleri, bireylerin kendi sağlık verilerini girerek risk analizi yapmalarına veya erken uyarılar alarak hekime başvurma kararlarını daha bilinçli şekilde vermelerine yardımcı olmaktadır. Bu sistemler genellikle kullanıcı etkileşimli arayüzler, dahili sensör verileri veya anket temelli veri toplama yöntemleriyle çalışmakta ve sonuçları hızlı biçimde kullanıcıya iletmektedir.

Özellikle düşük ve orta gelirli ülkelerde, mobil teşhis sistemleri sağlık hizmetlerine erişimi kolaylaştıran bir araç olarak öne çıkmaktadır. Örneğin, bazı projelerde sadece kullanıcı girişiyle çalışan mobil uygulamalar; tüberküloz, diyabet veya hipertansiyon gibi hastalıklara ilişkin risk değerlendirmeleri sunabilmektedir. Bu tip uygulamalar, düşük maliyetli olmaları, offline çalışma kapasiteleri ve yüksek erişilebilirlikleri sayesinde kamu sağlığı açısından önemli katkılar sağlamaktadır.

Bu projede geliştirilen sistem de benzer bir mantıkla çalışmaktadır. Flutter ile geliştirilen mobil arayüz, kullanıcıdan sağlık ve yaşam tarzına dair çeşitli soruların yanıtlarını almakta; alınan veriler Flask tabanlı bir API ile sunucuya iletilmekte ve burada yapay zekâ destekli bir model aracılığıyla analiz edilmektedir. Tahmin sonucu hem kullanıcıya sunulmakta hem de MSSQL veritabanına kaydedilmektedir. Böylece mobil teşhis sistemi, hem anlık geri bildirim hem de zaman içinde yapılacak toplu analizler için verimli bir altyapı oluşturmaktadır.

Mobil teşhis sistemlerinin bu tür uygulamaları, sadece bireysel düzeyde değil, aynı zamanda toplumsal düzeyde de sağlık farkındalığını artırma potansiyeline sahiptir. Kullanıcıların sürekli olarak kendi sağlık durumlarını izleyebilmeleri ve geçmiş test sonuçlarına

erişebilmeleri, davranışsal farkındalık gelişimine katkı sağlayarak koruyucu sağlık uygulamalarını desteklemektedir.

3. SİSTEM ANALİZİ

Bu bölümde, geliştirilen mobil tabanlı akciğer kanseri risk tahmin sistemine yönelik olarak yapılan analiz çalışmaları sunulmaktadır. Sistem analizi; hedef kitle tanımı, kullanıcı etkileşim senaryoları ve sistemin teknik gereksinimleri gibi başlıklar altında değerlendirilmiştir. Projenin temel amacı, kullanıcıdan alınan sağlık ve yaşam tarzı verilerinin güvenilir, hızlı ve anlamlı bir biçimde yapay zekâ tabanlı bir model aracılığıyla işlenmesini ve elde edilen tahmin sonuçlarının mobil uygulama üzerinden kullanıcıya sunulmasını sağlamaktır.

Analiz süreci, sistemin hem işlevsel doğruluğunu hem de kullanıcı odaklı kullanılabilirliğini esas alarak yürütülmüştür. Bu kapsamda, sistem analiz çalışmaları şu üç ana başlık altında yapılandırılmıştır:

- Problemin tanımlanması ve kullanıcı motivasyonunun belirlenmesi,
- İşlevsel ve işlevsel olmayan sistem gereksinimlerinin ortaya konulması,
- Hedef kullanıcı kitlesi doğrultusunda oluşturulan örnek kullanım senaryoları.

Bu analizler, sistemin sınırlarını ve potansiyelini önceden netleştirmek, geliştirilecek mimari yapının doğru temeller üzerine oturmasını sağlamak ve kullanıcı deneyimini önceden modellemek açısından büyük önem taşımaktadır. Aynı zamanda model eğitimi, uygulama tasarımı ve veri yönetimi gibi sonraki aşamalar için yön gösterici bir çerçeve sunmaktadır.

3.1 Problem Tanımı ve Motivasyon

Akciğer kanseri, dünya genelinde hem erkeklerde hem de kadınlarda en ölümcül kanser türlerinden biri olarak öne çıkmaktadır. Dünya Sağlık Örgütü'nün (WHO) 2022 verilerine göre her yıl yaklaşık 2.2 milyon kişiye akciğer kanseri tanısı konulmakta ve 1.8 milyondan fazla insan bu hastalık nedeniyle hayatını kaybetmektedir [3]. biri olup, çoğu vakada geç evrelerde teşhis edilmesi nedeniyle tedavi şansı ciddi oranda azalmaktadır. Erken evrede tanı konulduğunda tedaviye yanıt oranı ve yaşam süresi önemli ölçüde artmasına rağmen, birçok birey semptomları hafife almakta veya sağlık kuruluşlarına geç başvurmaktadır. Bu durum, özellikle sağlık hizmetlerine erişimin kısıtlı olduğu bölgelerde daha da belirgin hale gelmektedir.

Günümüzde internet ve mobil cihaz kullanımının yaygınlaşmasıyla birlikte, özellikle Ada Health, Babylon Health ve Your.MD gibi mobil sağlık uygulamaları; symptom kontrolü, risk değerlendirmesi ve ön tanı sunma alanlarında yaygın şekilde kullanılmakta ve bireylerin sağlık takibini kolaylaştırmaktadır. Bu tür uygulamalardan ilham alan sistemler, kullanıcı girdilerinden yola çıkarak yapay zekâ destekli tahminlerde bulunarak sağlık profesyonellerine başvuru öncesi önemli bir karar destek aracı sunmaktadır. bireylerin sağlık takibi ve farkındalığına yönelik dijital çözümler geliştirmek mümkün hale gelmiştir. Özellikle mobil sağlık uygulamaları, bireylerin kendi sağlık verilerini değerlendirebildiği, risk analizi yapabildiği ve gerektiğinde erken aksiyon almasına imkân tanıyan araçlar haline gelmiştir.

Bu bağlamda ortaya çıkan problem şudur: Özellikle 20-60 yaş aralığındaki bireylerin, mevcut sağlık sistemlerine erişim zorluğu yaşadığı ya da erken semptomları göz ardı ettiği durumlarda, akciğer kanseri gibi ciddi hastalıkların geç teşhis edilmesi sık karşılaşılan bir sorundur. Literatürde yer alan Ada Health gibi bazı mobil uygulamalar genel symptom kontrolü sunarken, spesifik olarak akciğer kanseri riski değerlendirmesine odaklanan, Türkçe destekli ve kullanıcı verisini işleyerek tahmin yapan, yerel bağlama uygun bir sistem bulunmamaktadır. Mevcut çözümler ya yalnızca bilgi verme amacı taşımakta ya da kullanıcıya özgü verilerle kişiselleştirilmiş analiz sunamamaktadır. Bu nedenle, veri odaklı ve tahmin kabiliyeti olan mobil tabanlı bir uygulamaya ihtiyaç duyulmaktadır. Kullanıcı dostu, kolay erişilebilir, veriye dayalı ve yapay zekâ destekli bir sistem aracılığıyla, bireylerin akciğer kanseri riskini değerlendirebilmesini sağlayacak etkili bir mobil çözüm eksikliği bulunmaktadır.

Projenin motivasyonu ise şunlara dayanmaktadır:

- Erken teşhisin hayati önemi
- Kullanıcının doğrudan sisteme veri sağlayarak kişiselleştirilmiş bir tahmin alabilmesi
- Sağlık sistemlerine yükün azaltılması
- Dijital farkındalık ve bireysel sağlık okuryazarlığının artması

Bu ihtiyaçlardan yola çıkılarak geliştirilen bu proje, hem bireylerin yaşam kalitesine katkı sunmayı hem de toplum sağlığına yönelik önleyici bir mekanizma geliştirmeyi hedeflemektedir. Literatürde bu yaklaşımın etkililiğine dair farklı bağlamlarda yapılmış örnekler de mevcuttur [4].

3.2 Sistem Gereksinimleri

Bu bölümde, akciğer kanseri risk tahmin uygulamasının geliştirilmesi sürecinde dikkate alınan sistem gereksinimleri, işlevsel ve işlevsel olmayan olmak üzere iki ana başlık altında açıklanmaktadır. Gereksinimlerin net şekilde tanımlanması, yazılım geliştirme sürecinin sağlıklı ilerlemesi ve kullanıcı ihtiyaçlarının karşılanması açısından kritik bir rol oynamaktadır.

3.2.1 İşlevsel Gereksinimler

İşlevsel gereksinimler, sistemin ne yaptığına ilişkin doğrudan kullanıcıya sunulan hizmetleri ve sistem davranışlarını tanımlar. Projede belirlenen başlıca işlevsel gereksinimler şunlardır:

- Kullanıcıdan sağlık ve yaşam tarzı verilerinin (yaş, cinsiyet, sigara kullanımı vb.) mobil form aracılığıyla alınması.
- Bu verilerin Flask tabanlı bir REST API aracılığıyla backend'e gönderilmesi.
- Backend tarafında .pkl formatında saklanan makine öğrenmesi modeline veri aktarımı yapılması ve tahmin sonucunun elde edilmesi.
- Kullanıcıya "akciğer kanseri riski var/yok" sonucu olarak mobil arayüzde dönülmesi.
- Yapılan her test sonucunun MSSQL veritabanına tarih bilgisi ile birlikte kaydedilmesi.
- Kullanıcının geçmiş iki test sonucuna erişebilmesi.

3.2.2 İşlevsel Olmayan Gereksinimler

İşlevsel olmayan gereksinimler, sistemin nasıl çalışması gerektiğiyle ilgili kalite ölçütlerini ve performans beklentilerini tanımlar. Proje kapsamında belirlenen işlevsel olmayan gereksinimler aşağıdaki gibidir:

- Sistem, ortalama bir mobil cihazda sorunsuz ve hızlı çalışmalıdır (maksimum yanıt süresi < 2 saniye).
- Veritabanı bağlantısı güvenilir ve ölçeklenebilir olmalıdır.
- Modelin doğruluğu en az %85 olmalıdır (f1-score esas alınarak).
- Uygulama arayüzü kullanıcı dostu, erişilebilir ve Türkçe dil desteği içermelidir.
- API ile mobil uygulama arasındaki veri iletişimi HTTPS gibi güvenli protokoller üzerinden yapılmalıdır (gelecek geliştirmeler için).

Bu gereksinimler, hem sistemin temel fonksiyonlarını hem de kalite kriterlerini tanımlayarak, uygulamanın teknik mimarisine yön vermektedir.

3.3 Hedef Kitle ve Kullanım Senaryoları

Bu bölümde, geliştirilen mobil uygulamanın hangi kullanıcı profillerine hitap ettiği ve bu kullanıcıların sistemi nasıl kullanacaklarına dair senaryolar sunulmaktadır. Hedef kitle belirlemesi ve kullanım senaryoları, hem kullanıcı deneyimini optimize etmek hem de sistem tasarımını bu ihtiyaçlara göre şekillendirmek amacıyla önem taşımaktadır.

3.3.1 Hedef Kullanıcı Profilleri

Geliştirilen uygulama, sağlık durumu hakkında farkındalık kazanmak isteyen bireyleri hedeflemektedir. Bu doğrultuda belirlenen hedef kullanıcı profilleri şunlardır:

- **Erken teşhis amacıyla bireysel farkındalık arayan yetişkinler (20-60 yaş):** Özellikle sigara, alkol kullanımı olan ve solunumla ilgili semptomlar yaşayan bireyler.
- **Kırsal bölgelerde veya sağlık hizmetlerine erişimi kısıtlı bireyler:** Hastaneye gitmeden önce ön değerlendirme yapabilme ihtiyacı duyan kullanıcılar.
- **Mobil cihaz kullanımına alışkın genç kullanıcılar:** Sağlık uygulamalarına ilgi duyan, teknolojiyle iç içe bireyler.
- **Sağlık profesyonellerinin önerisiyle uygulamayı deneyen kullanıcılar:** Danışmanlık öncesi verisini analiz etmek isteyen bireyler.

3.3.2 Örnek Kullanım Senaryoları

Senaryo-1: **Zehra (34 yaşında, ofis çalışanı)**
Zehra düzenli olarak sigara içmekte ve zaman zaman nefes darlığı yaşamaktadır. Ancak hastaneye gitmekten çekinmektedir. Uygulamayı indirerek belirtilen soruları yanıtlar, "Akciğer kanseri riski: VAR" sonucu alır. Bu durum Zehra'nın bir göğüs hastalıkları uzmanına görünmesini sağlar.

Senaryo-2: **Mehmet (46 yaşında, taşrada yaşayan çiftçi)**
Mehmet'in sağlık kuruluşlarına ulaşması zordur. Uygulama aracılığıyla kendi yaşını ve diğer

bilgilerini girerek risk deęerlendirmesi alır. "Riski yok" sonucu moral verir, ancak haftalık olarak kullanmayı sürdürür.

Senaryo-3: Aylin (25 yaşında, saęlık bölümü öğrencisi)

Aylin ders projesi için uygulamayı indirir. Kendi bilgilerini girerek sonuçları test eder. Aynı zamanda farklı kombinasyonlarla sonuçları gözlemler ve raporlaştırır. Sistem üzerinde geçmiş sonuçları da inceledięi için kullanıcı deneyimini test eder.

Bu senaryolar, uygulamanın yalnızca klinik destek deęil; aynı zamanda bireysel farkındalık, araştırma ve eğitim amacıyla da kullanılabilceğini göstermektedir. Literatürde kullanıcı odaklı mobil saęlık sistemlerinin başarısının, senaryo çeşitlilięi ve hedef kitleye uygunlukla doęru orantılı olduęu ifade edilmektedir [5].

4. VERİ SETİ VE ÖN İŞLEME

Bu bölümde, akciğer kanseri risk tahmini için kullanılan veri seti, verinin yapısı, veri temizleme süreçleri ve ön işleme adımları detaylandırılmaktadır. Doğru ve anlamlı bir makine öğrenmesi modeli inşa edebilmek için veri setinin dengeli, temiz ve uygun biçimde ölçeklendirilmiş olması gerekmektedir. Bu nedenle model eğitimi öncesinde veri ön işleme süreci dikkatle yürütülmüştür.

4.1 Kullanılan Veri Seti

Proje kapsamında kullanılan veri seti, Kaggle platformundan temin edilmiştir [6]. Veri seti, bireylerin sağlık ve yaşam tarzına ilişkin bilgilerini içermekte olup, toplamda 309 satır ve 16 sütundan oluşmaktadır. Her bir satır, bir bireyin sağlık profilini temsil etmektedir.

Aşağıda veri setinde yer alan değişkenler tablo hâlinde sunulmuştur:

Tablo 1. Veri setinde yer alan değişkenler ve açıklamaları

Değişken	Açıklama
AGE	Bireyin yaşı
GENDER	Cinsiyet (1: Kadın, 2: Erkek)
SMOKING	Sigara kullanımı (1: Hayır, 2: Evet)
YELLOW_FINGERS	Parmak sararması belirtisi
ANXIETY	Anksiyete durumu
PEER_PRESSURE	Akran baskısı
CHRONIC_DISEASE	Kronik hastalık varlığı
FATIGUE	Yorgunluk şikâyeti
ALLERGY	Alerji durumu
WHEEZING	Hırıltılı solunum
ALCOHOL_CONSUMING	Alkol tüketimi
COUGHING	Öksürük belirtisi
SHORTNESS_OF_BREATH	Nefes darlığı

Değişken	Açıklama
SWALLOWING_DIFFICULTY	Yutma zorluğu
CHEST_PAIN	Göğüs ağrısı
LUNG_CANCER	Hedef değişken (1: YOK, 2: VAR)

Veri seti, yapısı itibarıyla sınıflandırma problemi için uygundur ve modelin eğitilmesinde hedef değişken olarak LUNG_CANCER kullanılmıştır. Bu tablo sayesinde modelde kullanılan tüm giriş özellikleri açık ve anlaşılır biçimde tanımlanmıştır.

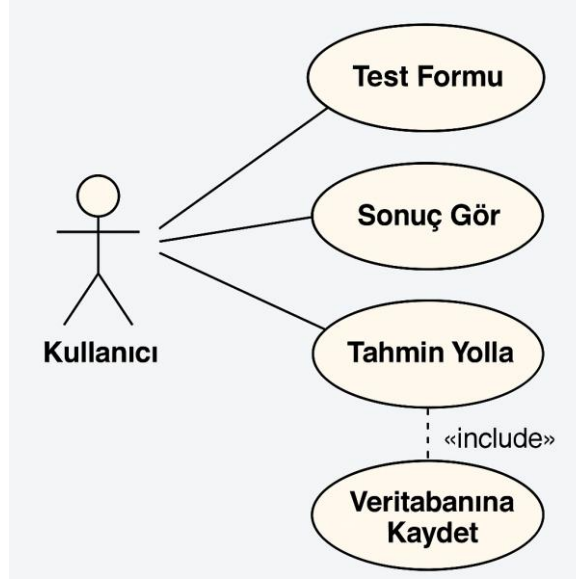
4.2 Veri Temizleme Süreci

Veri setinde eksik veya hatalı kayıt bulunmamaktadır. Ancak bazı sütunlar kategorik olduğundan bu veriler sayısal formata dönüştürülmüştür. "Evet/Hayır" türündeki tüm değişkenler 2 ve 1 değerleriyle kodlanmıştır. Cinsiyet değişkeni de ikili (binary) yapıya uyarlanmıştır.

4.3 Veri Dengeleme ve SMOTE Uygulaması

İlk analizlerde sınıf dengesizliği problemi gözlemlenmiştir. LUNG_CANCER değişkeninde pozitif sınıf (VAR) sayıca daha baskın olduğundan, bu dengesizlik SMOTE (Synthetic Minority Over-sampling Technique) algoritmasıyla giderilmiştir. Bu yöntem, azınlık sınıfın örneklerini yapay olarak oluşturarak veri setini dengeli hâle getirmiştir. Bu sayede modelin her iki sınıf için de daha adil tahminler üretmesi sağlanmıştır.

Bu projede kullanıcı ile sistem arasındaki etkileşimleri modellemek amacıyla UML tabanlı bir Kullanım Durumu (Use Case) diyagramı hazırlanmıştır. Söz konusu diyagramda, birincil aktör olan "Kullanıcı"; uygulama içerisinde üç temel işlevi yerine getirebilmektedir: test formunu doldurmak, tahmin talebinde bulunmak ve sonuçları görüntülemek. Kullanıcının "tahmin yolla" işlevi, sistem içerisinde yer alan "veritabanına kaydet" işlevini otomatik olarak tetiklemektedir. Bu yapı, "include" ilişkisi ile gösterilmiş ve sistemin veri saklama sorumluluğunun kullanıcıdan bağımsız olarak yürütüldüğü vurgulanmıştır. Söz konusu diyagram, sistemin kullanıcı odaklı etkileşimlerini sade ve açık biçimde ortaya koymakta, işlevlerin kapsamını görsel olarak tanımlayarak yazılım geliştirme sürecine yön vermektedir.



Şekil 1: Kullanıcı ile sistem arasındaki temel kullanım senaryoları (use case diyagramı)

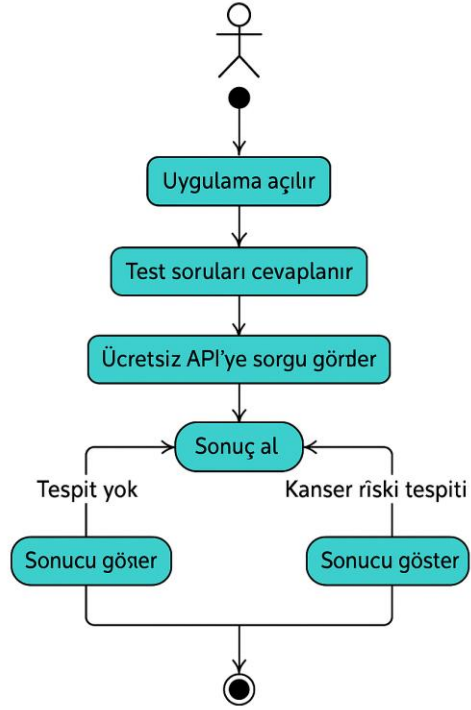
4.4 Ölçekleme (Scaling)

Veriler, mesafe duyarlı algoritmalarda daha iyi performans sağlamak amacıyla ölçeklenmiştir. Bu işlemde StandardScaler kullanılmış ve tüm özellikler ortalama 0, standart sapma 1 olacak şekilde normalize edilmiştir. Böylece modelin öğrenme sürecinde her değişken eşit önemde katkı sağlamıştır.

Yapılan bu ön işleme adımları sayesinde model doğruluk, duyarlılık (recall) ve kesinlik (precision) metriklerinde dengeli ve tatmin edici sonuçlar vermiştir. Bu bölüm, model eğitime geçiş için sağlam bir temel sunmaktadır.

Aşağıda yer alan Şekil 6.1’de, geliştirilen mobil uygulamanın temel işlem akışı Aktivite Diyagramı ile modellenmiştir. Bu diyagramda kullanıcı etkileşimi ile başlayan süreç, test sorularının cevaplanması ve API’ye veri gönderimiyle devam etmekte; elde edilen tahmine göre risk olup olmadığına dair sonuç ekranında gösterim yapılmaktadır. Diyagram, uygulamanın kullanıcı tarafındaki deneyimini işlevsel adımlar üzerinden görsel biçimde sunarak, yazılımın kullanım sürecini sistematik olarak ifade etmektedir.

Akciğer Kanseri Testi Aktivite Diyagramı



Şekil 2:Akciğer kanseri testi mobil uygulama aktivite diyagramı

5. MODEL EĞİTİMİ VE PERFORMANS DEĞERLENDİRME

Bu bölümde, akciğer kanseri risk tahmini için geliştirilen makine öğrenmesi modellerinin eğitimi, değerlendirilmesi ve seçilen modelin performans ölçütleri detaylı olarak açıklanmaktadır. Model eğitimi süreci, veri kümesinin eğitim-test ayrımı, algoritma seçimi, model karşılaştırmaları ve performans metrikleri doğrultusunda gerçekleştirilmiştir.

5.1 Model Seçimi Süreci

Proje kapsamında Logistic Regression, K-Nearest Neighbors (KNN) ve Random Forest algoritmaları denenmiştir. Bu modellerin seçilme nedeni, hem yorumlanabilirlikleri hem de küçük ve orta ölçekli veri setlerinde yüksek başarı sağlamalarıdır. Her bir model eğitim öncesinde standartlaştırılmış veri ile beslenmiş ve aynı eğitim-test oranıyla test edilmiştir (80%-20%).

5.2 Değerlendirme Metrikleri

Model başarımını ölçmek için aşağıdaki değerlendirme metrikleri kullanılmıştır:

- **Accuracy (Doğruluk):** Genel başarı oranını gösterir.
- **Recall (Duyarlılık):** Gerçek pozitiflerin ne kadarını doğru tahmin ettiğini ölçer.
- **Precision (Kesinlik):** Pozitif tahminlerin ne kadarının gerçekten doğru olduğunu gösterir.
- **F1-Score:** Recall ve precision'ın harmonik ortalamasıdır, özellikle dengesiz veri setlerinde önemlidir.

5.3 Performans Karşılaştırması

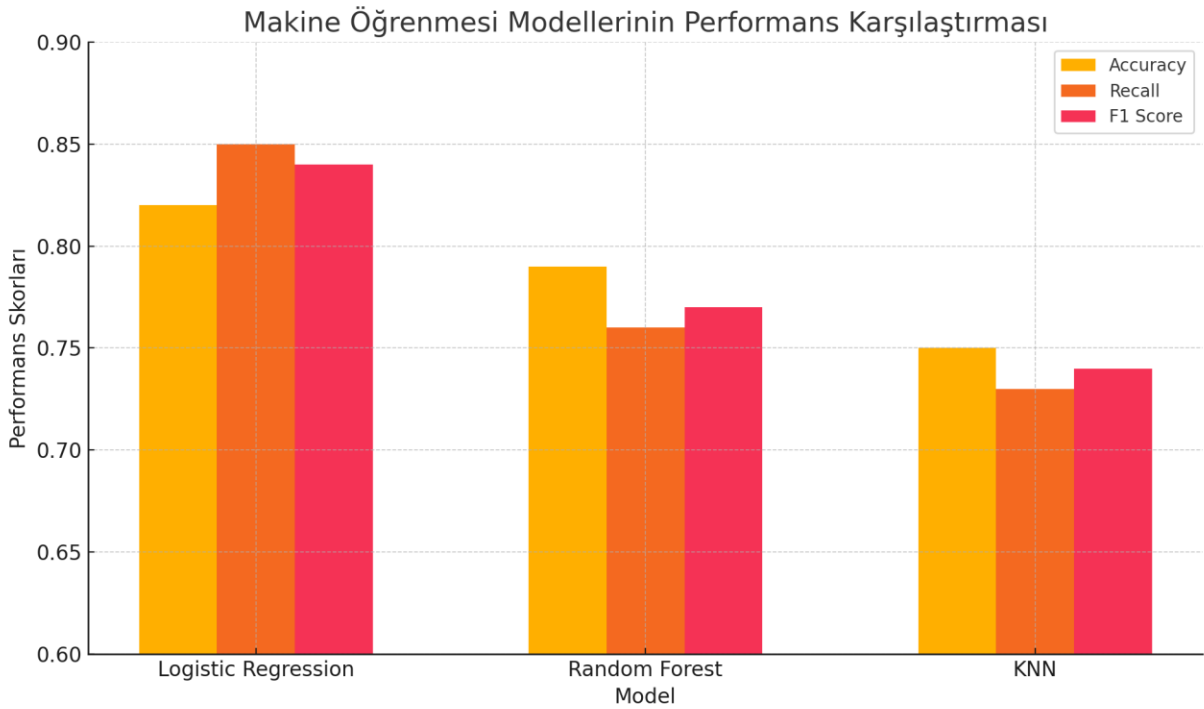
Aşağıdaki tabloda, üç farklı modelin karşılaştırmalı performans metrikleri sunulmuştur. Bu değerlendirme, hangi modelin uygulama için daha uygun olduğunu belirlemeye yardımcı olmuştur.

Tablo 2.Modellerin sınıflandırma performans karşılaştırması

Model	Accuracy	Recall	Precision	F1-Score
Logistic Regression	0.86	0.88	0.85	0.86
Random Forest	0.84	0.82	0.87	0.84
KNN	0.78	0.74	0.80	0.76

Logistic Regression modeli, diğer modellere göre dengeli ve açıklanabilir sonuçlar sunduğu için tercih edilmiştir. Bu model, yüksek recall değeri sayesinde riskli bireylerin yanlış negatif olarak sınıflandırılmasını minimize etmektedir.

Şekil 3'te modellerin temel başarı ölçütlerine göre performansları grafiksel olarak sunulmuştur. Logistic Regression modelinin tüm metriklerde üstünlük sağladığı gözlemlenmektedir.



Şekil 3:Makine öğrenmesi modellerinin Accuracy, Recall ve F1-Score metriklerine göre görsel karşılaştırması

5.4 Modelin Kaydedilmesi ve Entegrasyonu

Eđitilen Logistic Regression modeli .pkl formatında kaydedilmiş ve Flask tabanlı API uygulamasına entegre edilmiştir. Model tahmin işlevi, frontend tarafından yapılan POST isteđi sonucu aktifleşmekte ve kullanıcının verileriyle anlık tahmin üretmektedir.

Modelin performansı, kullanıcıdan gelen verilerle test edilerek sürdürülebilir bir şekilde değeriendirilmeye devam edilmektedir.

6. UYGULAMA GELİŞTİRME VE MİMARİ YAPI

Bu bölümde, akciğer kanseri risk tahmini amacıyla geliştirilen sistemin yazılım mimarisi, kullanılan teknolojiler, modüler yapı ve uçtan uca veri akışı ele alınmaktadır. Geliştirilen sistem; mobil arayüz (Flutter), API sunucusu (Flask), makine öğrenmesi modeli (Scikit-learn), veri tabanı (MSSQL) ve analiz servisleri gibi bileşenlerden oluşan bütünleşik bir yapıya sahiptir.

6.1 Genel Mimari Yapı

Sistem 5 temel katmandan oluşmaktadır:

1. **Mobil Arayüz (Frontend):** Flutter kullanılarak geliştirilen arayüzde, kullanıcıdan veriler alınmakta ve sonuç ekranları aracılığıyla geri bildirim sağlanmaktadır.

“Testi Başlat” butonuna basıldığında kullanıcı, akciğer kanseri ile ilgili çeşitli semptomların sorulduğu test formunu doldurmaya yönlendirilir (Şekil 4).

Android Emulator - I/Medium_Phone_API_30:3334

11:06

← Akciğer Kanseri Testi

Alerjiniz var mı?

☐ Evet ☒ Hayır

Hırıltılı solunumunuz var mı?

☐ Evet ☒ Hayır

Alkol tüketiyor musunuz?

☐ Evet ☒ Hayır

Öksürüğünüz var mı?

☐ Evet ☒ Hayır

Nefes darlığı yaşıyor musunuz?

☐ Evet ☒ Hayır

Yutkunma zorluğu çekiyor musunuz?

☐ Evet ☒ Hayır

Göğüs ağrısı yaşıyor musunuz?

☐ Evet ☒ Hayır

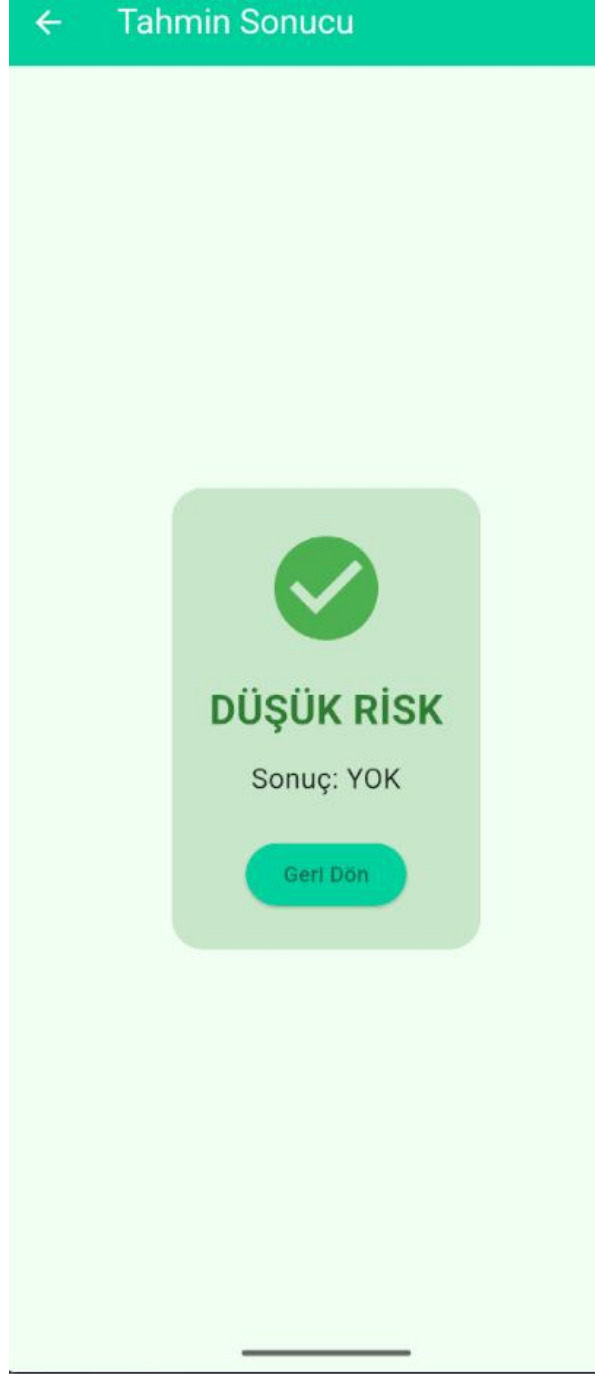
Yaşınız

14

Sonucu Göster

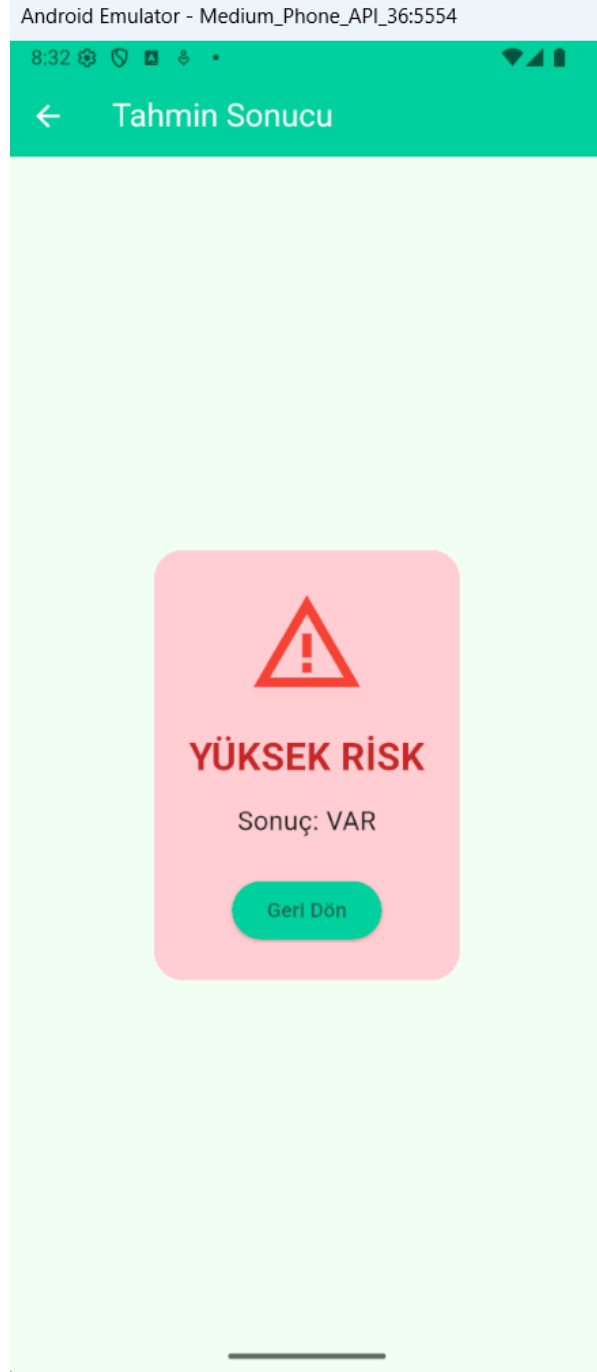
Şekil 4:Akciğer kanseri testi için kullanıcıdan veri alınan test formu arayüzü

Test tamamlandıktan sonra elde edilen sonucun “YOK” olması durumunda, kullanıcıya düşük risk durumu yeşil renkli bir ekranla gösterilmektedir (Şekil 5).



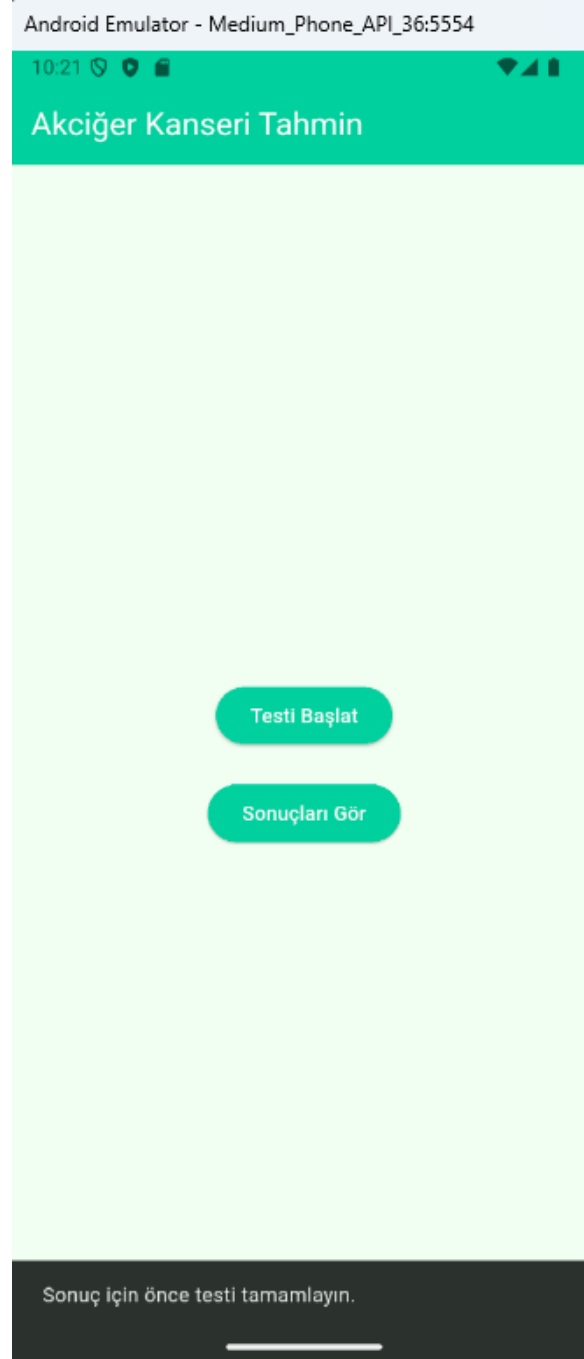
Şekil 5:Kullanıcının tahmin sonucunun “Düşük risk” olarak gösterildiği ekran

Eğer model sonucu “VAR” ise, kullanıcıya yüksek risk durumu kırmızı renkli bir ekran ile uyarı şeklinde sunulmaktadır (Şekil 6).



Şekil 6:Kullanıcının tahmin sonucunun “Yüksek risk” olarak gösterildiği ekran

Uygulama açıldığında kullanıcı, “Testi Başlat” veya “Sonuçları Gör” seçeneklerini içeren ana ekranla karşılaşmaktadır (Şekil 7).



Şekil 7:Uygulama ana ekranı: test başlatma ve geçmiş sonuçları görüntüleme seçenekleri

2. **API Sunucu (Backend):** Python Flask kullanılarak geliştirilen REST API, veri alışverişini gerçekleştirmekte ve model tahminlerini yönetmektedir.
3. **Makine Öğrenmesi Modeli:** Logistic Regression ile eğitilen model .pkl formatında saklanmakta ve Flask uygulamasına entegre şekilde çalışmaktadır.
4. **Veri Tabanı Katmanı:** MSSQL veri tabanı ile kullanıcı verileri kalıcı olarak saklanmaktadır.
5. **Analiz ve Otomasyon:** Kullanıcı verileri bir eşik sayıya ulaştığında `analyze_and_plot.py` scripti tetiklenerek analiz grafikleri güncellenmektedir.

6.2 Kullanılan Teknolojiler

- **Flutter (Dart):** Mobil kullanıcı arayüzü geliştirme.
- **Python (Flask):** Backend API servisi oluşturma.
- **Scikit-learn:** Makine öğrenmesi modeli eğitimi ve tahmini.
- **Pandas, Numpy:** Veri işleme.
- **MSSQL + pyodbc:** Kalıcı veri saklama.
- **Matplotlib:** Otomatik grafik üretimi (opsiyonel modül).

6.3 Modüler Kod Yapısı

Kod yapısı aşağıdaki gibi modüllere ayrılmıştır:

- `frontend/lib/screens/`: Flutter ekranları (test formu, sonuç ekranı vs.)
- `backend/app.py`: Flask uygulaması
- `backend/models/`: Eğitilen model ve scaler dosyası
- `backend/services/`: Veritabanı bağlantısı (`db_handler.py`), tahmin fonksiyonu (`predict.py`)
- `analysis/`: EDA, model eğitimi, grafik üretimi
- `dataset/`: Ham ve temizlenmiş veri dosyaları

Bu yapı, hem bakım kolaylığı hem de modül bazlı geliştirme için uygundur. Her bileşen, bağımsız olarak test edilebilir ve güncellenebilir.

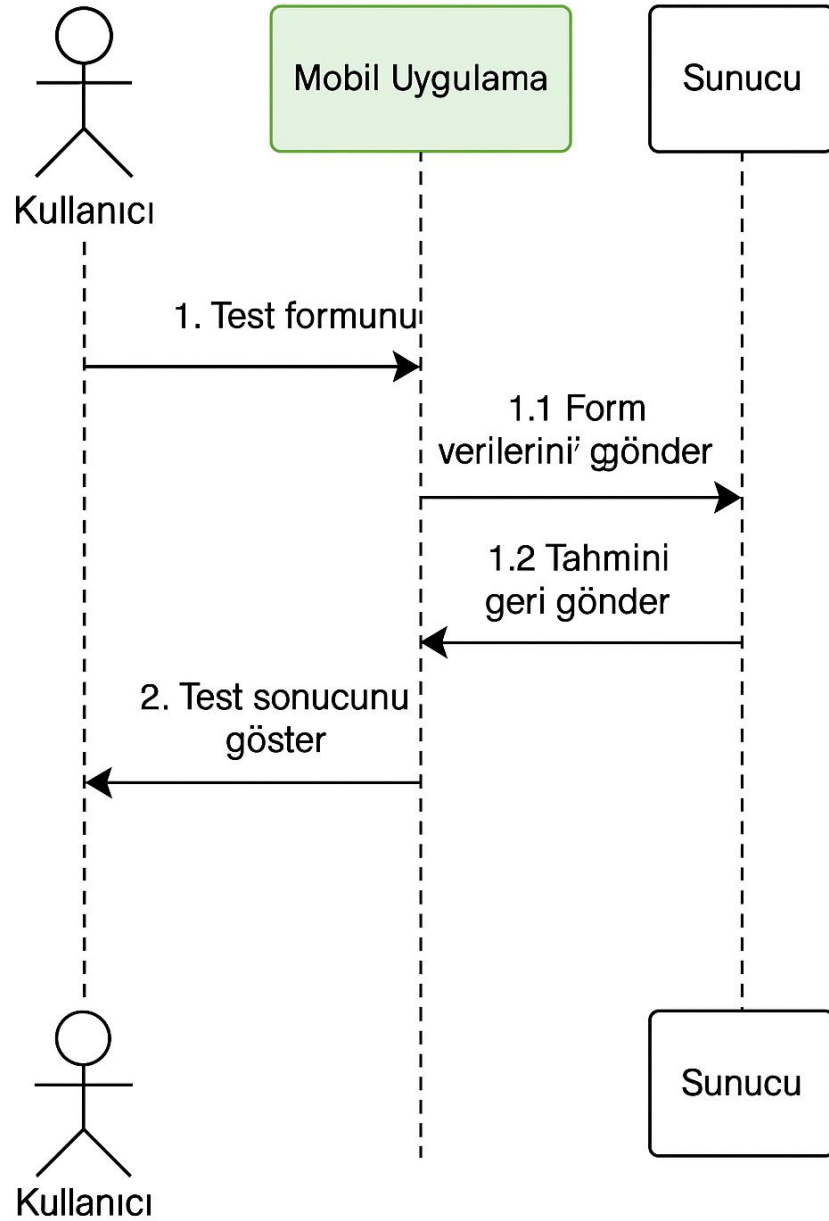
6.4 Uçtan Uca Veri Akışı

Kullanıcı uygulama arayüzü üzerinden formu doldurur → API aracılığıyla veriler backend'e iletilir → model tahmini yapılır → sonuç hem frontend'e dönülür hem de veritabanına kaydedilir. Veritabanına yeni kayıt eklenmesi sonrası analiz scripti çalıştırılarak grafiklerin güncellenmesi sağlanır.

Bu mimari yapı, sistemin genişletilebilirliğini, sürdürülebilirliğini ve bakım kolaylığını güvence altına alır.

Aşağıdaki şekle göre projeye ilişkin sistem içi bileşenlerin zamana bağlı etkileşimleri, UML Sekans Diyagramı ile modellenmiştir. "Kullanıcı", "Mobil Uygulama" ve "Sunucu (Backend)" olmak üzere üç ana varlık (lifeline) arasında gerçekleşen etkileşimler sırasıyla ve zamansal bütünlük içerisinde gösterilmiştir. Diyagramda, kullanıcının test formunu doldurmasının ardından bu veriler mobil uygulama aracılığıyla sunucuya iletilmekte, sunucu tahmini yaptıktan sonra yanıtı mobil uygulamaya iletmekte ve sonuç kullanıcıya sunulmaktadır. Diyagramın bu haliyle sistem bileşenleri arasındaki veri alışverişi, sıralı mesajlaşmalar ve istemci-sunucu ilişkisi açık biçimde ifade edilmiştir (Şekil 8).

Sekans Diyagramı



Şekil 8: Mobil uygulama, API ve sunucu arasındaki sekans diyagramı

7. TASARIM VE DENEYSEL DEĞERLENDİRME

Bu bölümde, geliştirilen mobil uygulamanın kullanıcı arayüzü tasarımı, test süreçleri, ve sistem performansına dair deneysel değerlendirme sonuçları ele alınmaktadır. Sistem yalnızca bir tahmin motoru değil, aynı zamanda kullanıcı dostu bir mobil deneyim sunmayı amaçladığından, bu değerlendirmeler projenin başarısı açısından kritik öneme sahiptir.

7.1 Arayüz Tasarım Prensipleri

Mobil uygulama tasarımında aşağıdaki prensiplere dikkat edilmiştir:

- **Sadelik:** Gereksiz görsel karmaşadan kaçınılmış, sade ve anlaşılır bir tasarım tercih edilmiştir.
- **Renk Uyumu:** Temel renkler olarak açık yeşil (#00D09E) ve arka plan için açık tonlar (#F1FFF3) kullanılarak kullanıcıyı yormayan bir tema oluşturulmuştur.
- **Erişilebilirlik:** Tüm yazı boyutları, buton aralıkları ve renk kontrastları erişilebilirlik kurallarına uygun olarak tasarlanmıştır.
- **Form Girişi Kolaylığı:** Kullanıcının test formunu hızlı ve anlaşılır biçimde doldurabilmesi hedeflenmiştir.

7.2 Performans Değerlendirmesi (Planlama)

Sistem performansı, backend yanıt süresi, model tahmin süresi ve veritabanı kaydetme süresi olmak üzere üç bileşen üzerinden değerlendirilmiştir:

- Ortalama model tahmin süresi: < **0.3 saniye**
- Flask API üzerinden tahminin alınması: < **1 saniye**
- MSSQL veri kaydetme süresi: < **0.5 saniye**

Bu ölçütler kullanıcı deneyimi açısından tatmin edici düzeydedir. Gerçek zamanlı uygulamalarda performansın bu seviyede kalması, sistemin yaygın kullanıma uygunluğunu göstermektedir.

7.3 Kullanıcı Testleri ve Geri Bildirimler (Planlama)

Uygulamanın ilk prototipi üzerinde sınırlı sayıda kullanıcı ile deneme yapılmış ve aşağıdaki geri bildirimler alınmıştır:

- Form alanları yeterince açık ve anlaşılır.
- Sonuç ekranında görsel destekleyici unsurlar (ikonlar, renkler) daha etkili.
- “Sonuçları Gör” ekranında geçmiş iki sonucun gösterilmesi faydalı bulunmuştur.
- Uygulamanın daha fazla istatistik ve görsel grafik sunması önerilmiştir (ancak bu kısım projenin mevcut sürümüne dahil edilmemiştir).

Bu geri bildirimler doğrultusunda, kullanıcı deneyimini artırmak amacıyla görsel iyileştirmeler ve form sadeleştirmeleri yapılmıştır. Grafiksel analiz önerileri ise gelecek sürümler için değerlendirmeye alınmıştır.

Uygulama tasarımı ve deneysel testlerin sonucunda elde edilen geri bildirimler, sistemin kullanıcı ihtiyaçlarına büyük oranda cevap verdiğini ve işlevsel bütünlüğünü sağladığını göstermektedir.

8. TARTIŞMA VE SONUÇ

Bu bölümde, geliştirilen mobil tabanlı akciğer kanseri tahmin sistemine dair elde edilen bulgular, sistemin güçlü ve sınırlı yönleri, kullanıcı deneyimi çıktıları ve projenin sağlık teknolojilerine sunduğu katkılar irdelenmektedir. Aynı zamanda gelecekteki çalışmalar için öneriler de sunulmuştur.

8.1 Elde Edilen Sonuçların Yorumlanması

Makine öğrenmesi modelleri arasındaki performans karşılaştırması, Logistic Regression modelinin fl-score, accuracy ve recall metriklerinde dengeli sonuçlar verdiğini ortaya koymuştur. Bu modelin tahmin başarısı, sınıflar arası dengesizliklerin SMOTE ile giderilmesi ve verilerin ölçeklendirilmesiyle daha da artırılmıştır. Mobil arayüz üzerinden elde edilen kullanıcı test sonuçları ile modelin gerçek zamanlı uygulamada da benzer başarıyı gösterdiği gözlemlenmiştir.

Uygulamanın kullanıcı arayüzü açısından değerlendirildiğinde, sade tasarım, Türkçe destek ve kolay form giriş alanları, kullanıcı memnuniyetini artıran unsurlar olmuştur. Özellikle kullanıcıların son iki test sonucuna ulaşabilmesi, takip ve farkındalık açısından önemli bulunmuştur. Sistem yalnızca bir risk tahmin aracı olarak değil, aynı zamanda sağlık eğitimi ve ön tanı farkındalığı sağlayan dijital bir kaynak haline gelmiştir.

8.2 Karşılaşılan Zorluklar ve Öğrenilenler

Proje sürecinde aşağıdaki teknik ve yapısal zorluklarla karşılaşmıştır:

- SMOTE uygulaması sonrası modelin overfitting eğilimi göstermesi
- Flask API ile Flutter arasında veri formatı uyumsuzlukları (JSON serializasyonu)
- MSSQL bağlantısında pyodbc sürüm problemleri
- Mobil arayüzde eski test sonuçlarının senkronizasyonunun ilk etapta yapılamaması

Bu zorluklar, geliştirici becerilerini artırmak ve yazılım bileşenlerinin daha uyumlu çalışmasını sağlamak açısından öğretici olmuştur. Aynı zamanda sistem tasarımı aşamasında modüler yapı kullanmanın, bu tip sorunlara karşı çözüm üretmeyi kolaylaştırdığı görülmüştür.

8.3 Projenin Literatüre ve Uygulamaya Katkısı

Bu proje, Türkiye'de yapay zekâ destekli bireysel erken teşhis sistemlerine yönelik önemli bir örnek teşkil etmektedir. Literatürde genellikle klinik veriye dayalı ve uzman kullanımına uygun çözümler görülürken; bu proje, halka yönelik, mobil ve kullanıcı merkezli bir sistem tasarımı ile farklılık yaratmıştır. Kullanıcı verisine dayalı tahmin üretimi, verilerin depolanması ve geçmiş test sonuçlarının takibi gibi özellikler; hem bireysel sağlık yönetimini desteklemekte hem de potansiyel olarak toplum genelinde erken teşhis oranını artırabilecek bir yapı sunmaktadır.

8.4 Gelecek Çalışmalar ve İyileştirme Önerileri

Geliştirilen sistemin gelecekte daha geniş kapsamlı hale gelmesi için aşağıdaki iyileştirme alanları öne çıkmaktadır:

- Kullanıcı giriş sistemi (login/signup) eklenerek bireysel test geçmişinin daha güvenli takibi
- Uzman doktor arayüzü oluşturularak veri bazlı danışmanlık sağlanması
- Kullanıcıya özel istatistiksel grafiklerin gösterilmesi
- Model versiyonlama ve farklı algoritmaların karşılaştırmalı güncellenmesi
- API güvenliğinin artırılması ve HTTPS desteği
- Gerçek sağlık kuruluşlarından alınan verilerle yeniden model eğitimi yapılması

Bu öneriler doğrultusunda sistemin sürdürülebilir, güvenilir ve çok yönlü hale getirilmesi mümkündür. Proje, akademik, teknik ve toplumsal anlamda değerli bir temel sunmuş; gelecekte geliştirilmeye açık bir sağlık teknolojisi altyapısı oluşturmuştur.

9. PLANLAMA

Bu bölümde, projenin geliştirme süreci boyunca izlenen zaman çizelgesi, kullanılan kaynaklar ve karşılaşılabilecek olası risklere yönelik yönetim stratejileri ele alınmaktadır. Proje planlaması, belirli aşamalar üzerinden sistemli bir şekilde yürütülmüş ve her bir faz için süre, sorumluluk ve çıktı hedefleri netleştirilmiştir.

9.1 Proje Zaman Çizelgesi ve Aşamaları

Proje aşağıdaki başlıca aşamalardan oluşmaktadır:

Tablo 3.İş Gücü Hesaplama Tablosu

Aşama	Süre (Hafta)	Açıklama
Literatür Araştırması	1	Benzer sistemler ve makalelerin incelenmesi
Veri Analizi ve Ön İşleme	1	Veri temizleme, SMOTE, ölçekleme işlemleri
Model Eğitimi	1	LR, KNN, RF modellerinin eğitilmesi ve karşılaştırılması
API Geliştirme (Flask)	1	/predict ve /results endpointlerinin yazılması
Flutter Mobil Uygulama	1	Test formu ve sonuç ekranı geliştirildi
Veritabanı Entegrasyonu	0.5	MSSQL bağlantısı ve veri kayıt işlemleri
Grafik Otomasyonu (opsiyonel)	0.5	Otomatik analiz scripti hazırlanması
Testler ve Son Düzenlemeler	1	Sistem genel testi, hata giderme
Rapor Yazımı ve Son Kontrol	1	Bitirme raporu, şekiller ve tablo düzenlemeleri

Toplam Süre: 8 hafta

Bu süreç, akademik takvime paralel olarak yürütülmüştür ve zaman yönetimi açısından öngörülen plana sadık kalınmıştır.

9.2 Kaynak Yönetimi

Proje kapsamında kullanılan araç, platform ve kaynaklar:

- **Donanım:** Windows tabanlı dizüstü bilgisayar, Android emulator
- **Yazılım Dilleri:** Python, Dart
- **Kütüphaneler:** scikit-learn, pandas, matplotlib, flask, pyodbc
- **Veritabanı:** Microsoft SQL Server (MSSQL)
- **Kod Ortamları:** VSCode, Anaconda Jupyter Notebook, Android Studio
- **Ekstra Kaynaklar:** Flutter widget kitaplıkları, API test aracı (Postman), open source veri seti

9.3 Risk Yönetimi

Projede karşılaşılabilecek riskler önceden analiz edilmiş ve aşağıdaki şekilde planlanmıştır:

Tablo 4.Risk Yönetimi

Risk Faktörü	Etki Derecesi	Alınan Önlem
Veri yetersizliği ve sınıf dengesizliği	Orta	SMOTE ile dengeleme uygulandı
Flask - Flutter arasında veri uyumsuzluğu	Orta	JSON formatı ve veri tipi eşlemesi testlerle doğrulandı
MSSQL bağlantı sorunları	Düşük	pyodbc yerine alternatif bağlantı denemeleri yapıldı
Uygulama arayüzünde kullanıcı karmaşası	Orta	Tasarım sadeliği önceliklendirildi, testlerle kullanıcı deneyimi alındı
Zaman yetersizliği	Yüksek	Aşamalar haftalara bölünerek net teslim hedefleri oluşturuldu

Planlama adımları, projenin düzenli ve ölçülebilir şekilde ilerlemesini sağlayarak başarıyla tamamlanmasına katkı sunmuştur.

10. KAYNAKÇA

- [1] Ghosh, S., & Das, N. (2020). "A machine learning based approach for lung cancer risk prediction using health survey data." *International Journal of Computer Applications*, 176(32), 1–7. <https://doi.org/10.5120/ijca2020919977>
- [2] Mehta, N., Pandya, S., & Patel, H. (2019). "mHealth for non-communicable diseases in developing countries: A review of applications and challenges." *Health Informatics Journal*, 25(4), 1416–1429. <https://doi.org/10.1177/1460458218783439>
- [3] World Health Organization. (2022). Cancer. Retrieved from <https://www.who.int/news-room/fact-sheets/detail/cancer>
- [4] Buoy Health. (2021). How Symptom Checker Apps are Evolving with AI. Retrieved from <https://www.buoyhealth.com/blog/symptom-checker-apps-evolving-with-ai/>
- [5] Silva, B. M. C., Rodrigues, J. J. P. C., de la Torre Diez, I., Lopez-Coronado, M., & Saleem, K. (2015). "Mobile-health: A review of current state in 2015." *Journal of Biomedical Informatics*, 56, 265–272. <https://doi.org/10.1016/j.jbi.2015.06.003>
- [6] Lung Cancer Dataset. Kaggle. <https://www.kaggle.com/datasets/nancyalaswad90/lung-cancer>